International Journal of Signal and Image Processing Issues
Vol. 2015, no. 1, pp. 9-17
ISSN: 2458-6498

Video Watermarking – Embedding binary watermark into the digital video using hybridization of three transforms

Ashish M. Kothari[1], Ved Vyas Dwivedi[2]

[1]Assistant Professor, Atmiya Institute of Technolgy & Science, Rajkot, Gujarat, India,
[2]Pro Vice Chancellor, C.U.Shah University, Wadhvan, Gujarat, India
Email: *amkothari.ec@gmail.com*

Abstract- This paper presents a unique method to embed binary image into the digital video stream. For this we made hybridization of three of the most powerful transforms in the domain of image processing namely Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD). For the sake of evaluation and comparison of the algorithm we used three parameters namely peak signal to noise ratio (PSNR), mean square error (MSE) and correlation.

Keywords: *Discrete Cosine Transform; Discrete Wavelet Transform; Singular Value Decomposition; copyright protection; robustness; digital video watermarking.*

I. INTRODUCTION

Today's era is the era of heralded connectivity which means that connectivity over internet and connectivity through wireless network. We do have some extraordinary inventions like digital camera, camcorders, MP3 players, PDA's etc. for creating, manipulating and enjoying the multimedia data. Nowadays the development of internet has given us some valuable gifts like electronic publishing of various files, e-advertising, e-newspaper, e-magazine, e-library, online video, online audio, on- line product ordering, online transactions, real time information delivery and many more. Because of all these storing, transmitting and distributing digital videos over the internet has been a very easy task. However creators of the videos are afraid of transmitting and distributing these valuable videos because of the problem of copyright protection. It is very easy task to copy digital data and when it is paste somewhere, it looks like the original one and therefore it leads towards the spiteful intent of what is called as piracy. The best possible way to protect multimedia data against illegal recording and retransmission is to embed a signal, called digital signature or copyright label or watermark into the cover medium that authenticates the owner of the data. The method is known as digital watermarking which a state is of art technique to put a secret message behind a cover medium in such a manner that the common man cannot visualize the message with a necked eye and he/she perceives it as a normal cover medium. Message may be the name of the creator, a logo of the company, or any other sign which can be extracted only when some specific algorithm is applied to extract the message and in this way the proof of ownership can be given. Nowadays the subject of interest is to provide proof of ownership and to prevent unauthorized tempering of the multimedia files. It can easily be done because editing of the files is done digitally. And this is the reason why both industry and academic people are working seriously on digital

watermarking. Techniques have been proposed for a variety of applications, including ownership protection, authentication and access control.

This research paper describes the digital watermarking where in we took video as a cover medium and grayscale image is chosen to be embedded in the video. The major factors that increases the demands of the Video watermarking [1, 2] are stated as below.

- Privacy of the digital data is required and because the copying of a video is comparatively very easy.
- Fighting against the "Intellectual property rights breach"
- Tempering of the digital video must be concealed.
- Copyright protection must not be eroded.

II. VISUAL QUALITY MATRICES

We have mainly used the following visual quality matrices [14] for the sake of comparison of degradation after the watermark is added to video.

$$\mathrm{MSE} = \frac{1}{\mathrm{M} \times \mathrm{N}} \sum_{x=1}^{M} \sum_{y=1}^{N} \{(f(x,y) - f'(x,y))\text{^}2\} \quad (1)$$

$$\mathrm{PSNR} = 10 \times \log \frac{255^2}{\mathrm{MSE}} \quad (2)$$

Here MSE – Mean Square Error

PSNR – Peak Signal to noise Ratio

f(x,y) – Original Frame of the video

f'(x,y) – Watermarked Frame of the Video.

The phrase peak signal-to-noise ratio [14], often abbreviated PSNR, is used to measure the similarity between two signals where in one is original and the other is altered version of the same. PSNR can be defined via the Mean Square Error (MSE) which gives us the idea of difference between the original and the altered signal. PSNR is measured in the logarithmic scale and MSE is measured in the general scale.

At the receiver end we extracted the watermark and measured the correlation [14] of the recovered watermark and original watermark for the sake of checking the robustness.

III. PROPOSED METHOD

A. Embedding Algorithm

Stepwise description of the embedding process of the proposed method is mentioned herewith.

1. Video is taken and converted into the sequence of frames.
2. First frame is taken and colorspace conversion is performed from RGB colorspace to the YCbCr colorspace. Here some of the reasons for choosing YCbCr colorspace are:
   - When JPEG compression is performed on the frame RGB colorspace is affected more as compare to the YCbCr colorspace. Here Y means luminance component of the frame and Cb and Cr are blue-difference and red-difference chrominance components.
   - It requires less disk space and less bandwidth.
   - This is the only colorspace used in the SD media, which have lower bandwidth and needs to have backwards compatibility.

Equations 3 to 5 show the conversion formulas from RGB to YCbCr colorspace while equations 6 to 8 show the reverse conversion.

$$Y = 16 + 65.481*R + 128.553*G + 24.966*B \quad (3)$$

$$Cb = 128 - 37.797*R - 74.203*G + 112*B \quad (4)$$

$$Cr = 128 + 112*R - 93.786*G - 18.214*B \quad (5)$$

$$R = Y + 0 * Cb + 1.402 * Cr \quad (6)$$

G = Y - 0.344136 * Cb - 0.714136 * Cr (7)

B = Y + 1.772 * Cb + 0 *Cr (8)

3. Y component of the frame is selected for the purpose of watermarking.

Here Y component is selected for the purpose of watermarking because human visual system is more sensitive towards change in the brightness compare to the color. Therefore JPEG compression does not compress any sample of the Y component. So if message content is embedded in the Y component it may be better preserved as compare to Cb and Cr. Also even if the frame is compressed a normal person perceives same information as the original case.

4. A two dimensional Discrete Cosine Transform [3, 4, 5, 6] is applied on the Y frame.

Equations 9 and 10 show the expressions of the 2D DCT and 2D inverse DCT respectively where f(x, y) is a spatial domain representation and F(u, v) is a frequency domain representation. Discrete cosine Transform converts a spatial domain 2D representation into its frequency domain equivalent. There are some of the observations of what appears at the output of DCT. One of the observations is that the number of rows and columns in the transformed matrix is exactly equal to that in the original 2D matrix. Another observation is that the upper left corner of the DCT transformed matrix contains most of the signal energy because the upper left corner contains the low frequency part of the original signal with the left most component is a DC coefficients and all the other signals are AC coefficients. Another observation is that going towards the right side in zigzag fashion the frequency increases and the values of the coefficients starts decreasing. Another observation is that the DC coefficient is always an integer and the range of that would be in between -1024 to 1023 while AC coefficient may be integer or non-integer.

$$F(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cos[\frac{(2x+1)u\pi}{2M}]\cos[\frac{(2y+1)v\pi}{2N}] \quad (9)$$

where $\alpha(u) = \sqrt{1/M}$ for *u=0;*

$\alpha(u) = \sqrt{2/M}$ for *u=1,2,3,.....M-1;*

$\alpha(v) = \sqrt{1/N}$ for *v=0;*

$\alpha(v) = \sqrt{2/N}$ for *v=1,2,3,.....N-1;*

$$f(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} \alpha(u)\alpha(v)F(u,v)\cos[\frac{(2x+1)u\pi}{2M}]\cos[\frac{(2y+1)v\pi}{2N}] \quad (10)$$

where *x = 0,1,2,……..M-1, y = 0,1,2,……..N-1*

5. A three level Discrete Wavelet Transform (DWT) [7, 8, 9 and 10] is applied on the DCT transformed frame.

Wavelets as compare to waves, which are oscillating functions of time or space and are periodic, are localized waves having their energy concentrated in time or space and they are used for the purpose of analyzing a signal. Wavelet transform convolves the signals with particular instances of wavelets at various time scales and positions. The DWT is based on sub-band coding, is easy to implement, does require limited time and resources and yields fast computations of wavelet transform. The DWT is a combined process of filtering and sub-sampling where sub-sampling may be up-sampling or down-sampling. Filtering operation determines the resolution, which is a measure of the information in the signal, of the signal while sub-sampling operation determines the scale of the signal. The multi-level DWT is computed using a successive low pass and high pass filtering. At each decomposition level, frequency resolution is doubled as the uncertainty in frequency is reduced by half and time resolution is made half means if the signal has originally of 500 samples it reduces to 250 samples at the end of first decomposition level. Thus it may be observed that with this approach, time resolution becomes good at high frequencies while frequency resolution becomes good at low frequencies.

6. Singular Value decomposition [15-19] is applied to both DWT transformed frame and the message frame.

Singular value decomposition (SVD) is a numerical technique based on the linear algebra and it is used to diagonalize matrices in numerical analysis. When SVD is applied to an Image A of size MxN, results are three matrices, namely U, V and S. Here U and V matrices are called unitary matrices having size MxM and NxN respectively and S matrix is called diagonal matrix having size MxN. SVD is able to efficiently represent the

intrinsic algebraic properties of an image, where singular values correspond to the brightness of the image and singular vectors reflect geometry characteristics of the image. The image A is represented as given in equation 11.

$$A = U * S * V^{T} \tag{11}$$

The columns of the U matrix are called the left singular vectors while the columns of the V matrix are called the right singular vectors of A. The diagonal entries of S are called the singular values of A and are arranged in decreasing order. The singular values (SVs) of an image have very good stability, i.e., when a small perturbation is added to an image, its SVs do not change significantly. An image matrix has many small singular values compared with the first singular value. Even ignoring these small singular values in the reconstruction of the image does not affect the quality of the reconstructed image.

7. The singular value of the frame is modified according to the singular values of message.
8. Inverse SVD is applied to get the watermarked DWT frame.
9. Inverse three level DWT is performed to get watermarked DCT frame.
10. Inverse DCT is applied to get the watermarked Y frame.
11. Inverse colorspace conversion is performed to get watermarked frame.
12. Next frame is taken and steps 2 to 9 are performed until all the frames are watermarked.
13. All watermarked frames are combined to get the watermarked video.

Figure 1 shows the results of the embedding process with the gain factor 100.

(a)

(b)

Figure 1: *Example of Invisible Video Watermarking using Hybrid Method with Gain Factor = 100 and binary message (a) First five frames of original video (b) Watermarked Frames*

B. Extraction Algorithm

Stepwise description of the extraction process of the invention is mentioned herewith.
Watermarked Video, which may possibly be attacked, is taken and converted into the sequence of frames.

1. First frame is taken and colorspace conversion is performed from RGB colorspace to the YCbCr colorspace.
2. Y component of the frame is selected for the purpose of watermark extraction.
3. A two dimensional Discrete Cosine Transform is applied on the Y frame.
4. A three level Discrete Wavelet Transform (DWT) is applied on the DCT transformed frame.
5. Singular Value decomposition is applied to DWT transformed frame.
6. Singular values are modified to get the watermark message back.

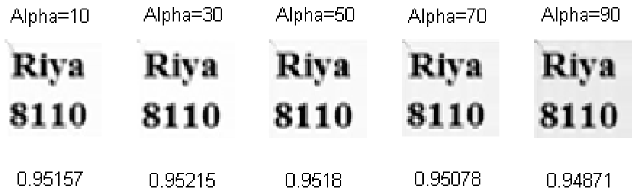
Figure 2 shows the result of the recovered message from the video.

Figure 2: *Recovered Messages*

C. Results

Figure 3 shows results of the method on frame 1 considering various values of the gain factor.

(a)



(b)

Figure 3: *Results with various gain factors (a) Watermarked Frame 1 (b) Recovered Messages*

Table 1 shows the results of the first five frames of the video taking gain factor to be 100. Table 2 shows the results of watermarking frame 1 with various gain factors.

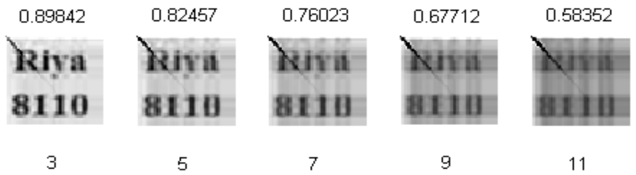
Frame No.	PSNR(dB)	MSE	Correlation
1	37.0173	12.9227	0.9332
2	37.0311	12.8818	0.9316
3	37.275	12.178	0.8449
4	37.3222	12.0464	0.928
5	37.239	12.2796	0.9401

TABLE 1. *PSNR Results of hybrid Method with gain factor = 100 and Binary message*

Alpha	PSNR(dB)	MSE	Correlation
1	Inf	0	0.8775
2	79.8438	0.0007	0.9069
3	72.6838	0.0035	0.9348
4	69.1053	0.008	0.9424
5	66.0303	0.0162	0.9465
6	63.7036	0.0277	0.949
7	61.6699	0.0443	0.9498
8	59.8345	0.0676	0.9508
9	58.1916	0.0986	0.9513
10	56.7222	0.1383	0.9516
100	37.0173	12.9227	0.9472

TABLE 2. *Results of hybrid Method on frame 1 using various gain factors with binary message*

Figure 4 shows results of the scheme with gain factor 100 and under various attacks. Here the numbers on the upper side of the frames show PSNR and that in the lower side of the frames show variants of particular attacks. Similarly the numbers on the upper side of the recovered watermark show correlation values.


0.89842
0.82457
0.76023
0.67712
0.58352
Riya 8110
3
5
7
9
11

(a)


36.0383
34.277
34.1943
34.1351
34.123
0.5
1.5
2
3
3.5
0.9363
0.9024
0.90047
0.89927
0.89902
Riya 8110
0.5
1.5
2
3
3.5

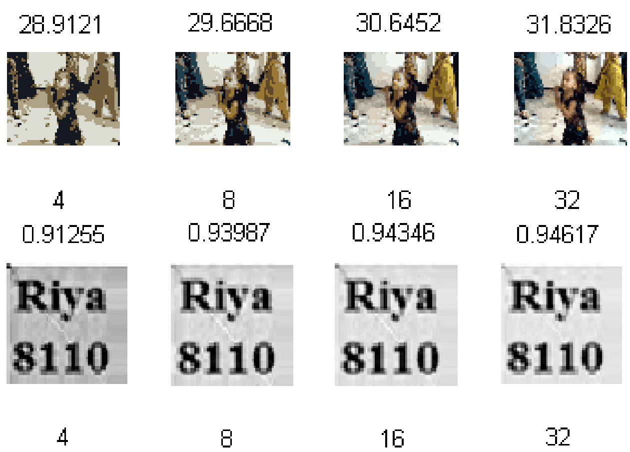
(b)


35.3738
33.9968
33.1277
32.5073
32.0223
3
5
7
9
11
0.92292
0.88365
0.84732
0.80021
0.72882
Riya 8110
3
5
7
9
11

(c)


30.4355
34.3707
35.9547
36.3589
36.4441
5
20
60
80
95
0.89322
0.92686
0.93788
0.94225
0.94595
Riya 8110
5
20
60
80
95

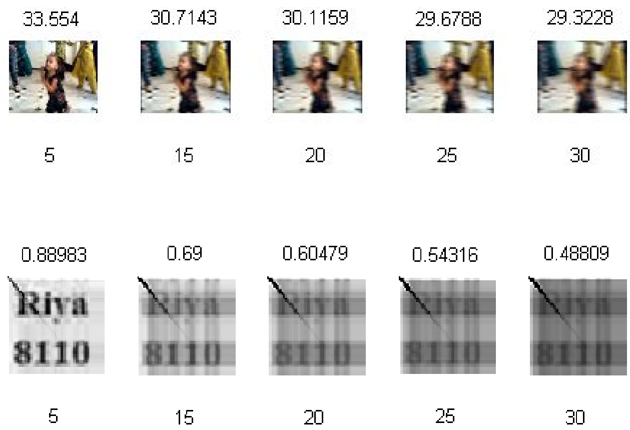
(d)


28.9121
29.6668
30.6452
31.8326
4
8
16
32
0.91255
0.93987
0.94346
0.94617
Riya 8110
4
8
16
32

(e)


28.7829
0.9433
Riya
8110

(f)


33.554 30.7143 30.1159 29.6788 29.3228
5 15 20 25 30
0.88983 0.69 0.60479 0.54316 0.48809
5 15 20 25 30

(g)


26.7284 26.854 26.7553 26.6996 26.7824
30 90 135 225 270
0.70167 0.39131 0.68811 0.73047 0.54739
30 90 135 225 270

(h)


34.0352 29.0648 27.9524 27.6127 27.6003
0.0005 0.01 0.09 0.8 1
0.94637 0.93865 0.85429 0.62432 0.60742
0.0005 0.01 0.09 0.8 1

(i)


37.0012 36.6888 34.6573 28.3132 27.4415
0.0005 0.01 0.09 0.8 1

Figure 4: *Results of Watermarked Frame and Recovered Message under various attacks (a) Average Filtering with various mask-size (b) Gaussian Law Pass Filtering with various standard deviations (c) Median Filtering with various mask-size (d) Compression with various Quality values (e) Color Reduction with various no. of colors (f) Histogram Equalization (g) Linear Motion of Camera with various no. of pixels (h) Rotation with various angles (i) Gaussian Noise with 0 Mean and various variances (j) Salt & Pepper Noise with various variances (k) Spackle Noise with various variances (l) Cropping with various crop region*

IV. CONCLUSION

Following are some of the observations made after successfully implemented both embedding and extracting algorithm. Here a gain factor of 100 is assumed for the sake of observations in terms of perceptibility and robustness. Higher is the value of PSNR, higher is the perceptibility and higher is the value of correlation, higher is the robustness.

- Perceptibility in this method decreases with the increase in gain factor.
- Robustness decreases with the increase in gain factor.
- The frames looks visibly fine if the resultant PSNR is above 28 dB. The message seems visibly identifiable if the resultant correlation is greater than 0.50.
- This method is fully robust against all kind of attacks.

Comparison with Spatial and Transform Domain Methods

- Perceptibility in this method is considerably highest among all methods at the same gain factor.
- Robustness achieved is considerably higher than Correlation, DCT and DWT based methods and slightly

less than SVD based method.

REFERENCES

[1] Paul, Rini T. Review of robust video watermarking techniques. IJCA Special Issue on Computational Science 2011 vol. 3, p. 90-95.

[2] CornelisLangelaar, Gerhard, and Reginald L. Lagendijk. Watermarking digital image and video data–A state of Art Review. IEEE SIGNAL PROCESSING MAGAZINE 2000.

[3] Kaur, Blossom, Amandeep Kaur, and Jasdeep Singh. Steganographic approach for hiding image in DCT domain. International Journal of Advances in Engineering & Technology 2011, vol. 1, issue 3 p. 72-78.

[4] Bhattacharyya, Debnath, Tai-hoon Kim, and Gang-soo Lee. Watermarking Using Multiresolution Cosine Transformation: A Review. Signal Processing, Image Processing and Pattern Recognition. Springer Berlin Heidelberg, 2011, P. 126-139.

[5] Hsu, Chiou-Ting, and Ja-Ling Wu. DCT-based watermarking for video. Consumer Electronics, IEEE Transactions on 44.1 1998, p. 206-216.

[6] Shoemaker, Chris. Hidden bits: A survey of techniques for digital watermarking. Independent study, EER 290, 2002.

[7] Xia, Xiang-Gen, Charles G. Boncelet, and Gonzalo R. Arce. A multiresolution watermark for digital images. Image Processing, 1997. Proceedings., International Conference on. Vol. 1. IEEE, 1997.

[8] Ejima, Masataka, and Akio Miyazaki. A wavelet-based watermarking for digital images and video. IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences 2000, vol. 83, issue 3, p. 532-540.

[9] Daren, Huang, et al. A DWT-based image watermarking algorithm. null. IEEE, 2001.

[10] Kothari, Ashish M., and Ved Vyas Dwivedi. Video Watermarking–Combination of Discrete Wavelet & Cosine Transform to Achieve Extra Robustness. International Journal of Image, Graphics and Signal Processing 2013 vol. 5, issue 3, p. 36.

[11] Kutter, Martin, and Fabien AP Petitcolas. Fair benchmark for image watermarking systems. Electronic Imaging'99. International Society for Optics and Photonics, 1999.

[12] Gonzalez, Rafael C., Richard Eugene Woods, and Steven L. Eddins. Digital image processing using MATLAB. Pearson Education India, 2004.

[13] Yushen, Li, Hao Yanling, and Wang Chenye. A research on the robust digital watermark of color radar images. Information and Automation , 2010 IEEE International Conference on. IEEE, 2010.

[14] Santhi, V., and Dr Arunkumar Thangavelu. DWT-SVD combined full band robust watermarking technique for color images in YUV color space. International Journal of Computer Theory and Engineering 2009 vol. 1, issue 4, p. 424-429.

[15] Andrews, Harry C., and Claude L. Patterson III. Singular value decomposition (SVD) image coding. Communications, IEEE Transactions on 1996, vol.24, issue 4, p. 425-432.

[16] Golub, Gene H., and Christian Reinsch. Singular value decomposition and least squares solutions. Numerische mathematik 1970, vol. 14, issue 5 p. 403-420.

[17] Liu, Ruizhen, and Tieniu Tan. An SVD-based watermarking scheme for protecting rightful ownership. Multimedia, IEEE Transactions on 4.1 2002, p. 121-128.

[18] Chang, Chin-Chen, Piyu Tsai, and Chia-Chen Lin. SVD-based digital image watermarking scheme. Pattern Recognition Letters 2005, vol. 26, issue 10, p. 1577-1586.

[19] Wu, Yongdong. On the security of an SVD-based ownership watermarking. IEEE transactions on Multimedia 2005, vol. 7, issue 4, p. 624-627.